

The turnpike theorems for Markov games ^{1 2}

Vassili Kolokoltsov and Wei Yang

Department of Statistics, University of Warwick

Coventry, CV4 7AL, UK

v.kolokoltsov@warwick.ac.uk wei.yang@warwick.ac.uk

Abstract

This paper has a two-folded purpose. First, we attempt to outline the development of the turnpike theorems in the the last several decades. Second, we study turnpike theorems in finite-horizon two-person zero-sum Markov games on a general Borel state space. Utilising the Bellman (or Shapley) operator defined for this game, we prove the stochastic versions of the early turnpike theorem on the set of optimal strategies and the middle turnpike theorem on the distribution of the state space.

Key Words and Phrases: Markov games, general Borel state space, turnpike theorem, average cost, turnpike on strategies, turnpike on the distribution of states, middle turnpike, early turnpike

1 Introduction

Turnpike theorems are a set of economic theorems on optimal control problems, which characterise solutions (optimal paths) of optimal control problems and focus on the stationary optimal one. The *turnpike theorem* was first proposed by Dorfman, Samuelson and Solow in [27] within the context of an optimal growth model. In the literature, there are extensive work on the turnpike theorems in many diversified settings of optimal control problems and games, ranging from discrete time to continuous time, from the deterministic case to the stochastic case, with or without discounting. There is no clear-cut and unified classification for turnpike theorems. For example, McKenzie [67] distinguished three kinds of turnpikes: the early, late and

¹Supported by the AFOSR grant FA9550-09-1-0664 'Nonlinear Markov control processes and games'

²Submitted to *Dynamic Games and Applications* for publication

middle turnpikes, but in [71], he used a different classification scheme with the *Samuelson* and *Ramsey turnpike*. We think that this is helpful to clarify different types of turnpike theorems and to give a clear picture on the state of this subject. Hence we start this paper with a survey on the main developments of the turnpike theorems in the last several decades.

Further on, in this paper, we study turnpike theorems in finite-horizon zero-sum Markov games on a general Borel state space following the approach developed in [50] in which a finite state space was considered. A general approach to Borel state space stochastic games was developed by Maitra and Sudderth [65]. They proved the existence of the game value for general universally measurable strategies. However, the optimal strategies need not exist in their games and it is assumed the transition probability function is strongly continuous in the actions of one player.

In our setting, under the weaker assumption (4.1) in Section 4, we show the convergence of the long time average cost (reward) of the game and prove the turnpike theorems on the set of strategies and on the distribution of the states.

The organisation of this paper is as follows. In Section 2, we sketch a picture of turnpike theory showing how the theory was developed. In Section 3, we present the basic setting for a two-person zero-sum Markov game and define the Bellman (or Shapley) operator. The main results will be presented and proved in Section 4. Section 5 concludes this paper and proposes further research.

2 Recent development in the turnpike theory

In this section, we outline the development of the turnpike theory in the last several decades. We shall start with turnpike theorems in the deterministic setting, followed by the stochastic analog of the turnpike theorems.

2.1 The turnpike theorems in the deterministic setting

For the turnpike theorems in the deterministic setting, it is important to distinguish the *asymptotic* and *neighborhood turnpike theorems* for solutions to *discounted infinite-horizon* optimal control problems from the *classical turnpike theorems* for solutions to *undiscounted finite-horizon* problems. Some surveys in the deterministic setting can be also found in [71, 94, 7, 48].

The *classical turnpike theorem* considers a deterministic finite-horizon ($T > 0$) optimal control problem on a given state space X . In a discrete time case, the problem is

$$\begin{cases} \text{maximise} & \sum_{t=0}^{T-1} L(x_t, x_{t+1}) \\ \text{subject to} & x_0, x_T \in X \text{ given, } \{x_t\}_{t=0}^T \subset X \end{cases} \quad (2.1)$$

with a cost (utility) function $L : X^2 \rightarrow R$, $T \in \mathbb{N}$. In a continuous time case, the problem is

$$\begin{cases} \text{maximise} & \int_0^T L(x_t, u_t) dt \\ \text{subject to} & \dot{x}_t = h(x_t, u_t), \quad u_t \in U \\ & x_0, x_T \in X \text{ given, } \{x_t\}_{t \in [0, T]} \subset X \end{cases} \quad (2.2)$$

with a cost (utility) function $L : X \times U \rightarrow R$, where $h : X \times U \rightarrow R$ is a continuous mapping with respect to both arguments, U is the set of admissible controls and $T \in \mathbb{R}^+$. The initial value x_0 and the terminal value x_T are held fixed. Let x^* denote a solution of the maximization problem

$$\text{maximise } L(x, x), \quad x \in X.$$

The path $x^* := \{x_t : x_t = x^*, \forall t \in [0, T]\}$ is called a stationary optimal path. Often, the uniqueness of the stationary optimal path is assumed.

In the case where the terminal value x_T is not given, an additional term, interpreted as the terminal payoff function, should be added to the above problems. That is, the optimal problem is, in a discrete time setting

$$\begin{cases} \text{maximise} & \sum_{t=0}^{T-1} L(x_t, x_{t+1}) + S(x_T) \\ \text{subject to} & x_0 \text{ given, } \{x_t\}_{t=0}^T \subset X \end{cases} \quad (2.3)$$

and, in a continuous time case

$$\begin{cases} \text{maximise} & \int_0^T L(x_t, u_t) dt + S(x_T) \\ \text{subject to} & \dot{x}_t = h(x_t, u_t), \quad u_t \in U \\ & x_0 \text{ given, } \{x_t\}_{t \in [0, T]} \subset X \end{cases} \quad (2.4)$$

where the terminal payoff function $S : X \rightarrow \mathbb{R}$.

The original formulation of the continuous time problem (2.2) is ascribed to Samuelson-Solow [82] and the discrete time problem (2.1) to Gale [32] and McKenzie [78, 67]. It was further investigated by e.g. Radner [80], Morishima [77], McKenzie [66] and Makarov and Rubinov [64], Atsumi [1] and Haurie [35]. McKenzie [71] discussed in detail the development of the turnpike theory in both the von Neumann growth model and the Ramsey's growth model and specified a classification of Samuelson turnpike and Ramsey turnpike. The classical techniques which were developed to prove the classical turnpike theorems are critically dependent on the convexity of the control set U and technology function h and the preference function L . These classical techniques fall into two classes: Value-Loss methods and Monotonicity method, see more discussion in [45].

The classical turnpike theorem states that if T is large enough, the optimal path $\{x_t^*\}$ that transfers the system from x_0 to x_T approaches the unique optimal stationary level x^* , and stays close to it for a large fraction of T , and moves away toward the terminal state only in the final periods; that is, for each $\epsilon > 0$, there exist an integer T_0 such that for each $T \geq 2T_0$,

$$\|x_t^* - x^*\| \leq \epsilon, \quad \forall t \in [T_0, T - T_0]. \quad (2.5)$$

This is called the *middle turnpike theorem* in [67]. Moreover, if this result holds for any initial value $x_0 \in X$, it is the *global* version; if the result holds for some initial value x_0 which is near enough to x^* , it is the *local* version, e.g. [9]. Further, if for each $\epsilon > 0$, there exists an integer T_0 and a number $\delta > 0$ such that $\|x_0 - x^*\| \leq \delta$ implies (2.5) for all $t \in [0, T - T_0]$ with $T \geq 2T_0$, this was called the *early turnpike theorem* in [67]. Since this image resembles a map of an intercity highway or turnpike with entrance and exit ramps, a stationary optimal path x^* is often referred to as a *turnpike*.

A version of the classical turnpike theorem for games was also presented. For example, two-player zero-sum differential games were studied by Zaslavski [91], showing that there exists a pair of optimal stationary paths x^* and y^* for the two players and presenting a version of classical middle turnpike theorem, and by Alvarez and Bardi [3, 2, 8], proving a turnpike theorem for the value of the game.

In the classical setting, the state space X was often a compact, convex, bounded and closed subset of a finite-dimensional space; the cost (utility) function L is often continuous, smooth and strictly concave (or convex); hence the turnpike theorem was proved by the concavity analysis. Later,

under upper (lower) semicontinuity, some extensions to general Banach state spaces were obtained by e.g. Kolokoltsov [50], Yakovenko and L.A. Kontorer [88], Zaslavski [90, 92, 93], Mamedov [59].

When the future is discounted, the deterministic optimal control problem with infinite-horizon is described, in a discrete-time setting as

$$\begin{cases} \text{maximise} & \sum_{t=0}^{\infty} \delta^t L(x_t, x_{t+1}) \\ \text{subject to} & x_0 \in X, \quad \{x_t\}_{t=0}^{\infty} \subset X \end{cases} \quad (2.6)$$

and in a continuous-time setting as

$$\begin{cases} \text{maximise} & \int_0^{\infty} e^{-\rho t} L(x_t, u_t) dt \\ \text{subject to} & \dot{x}_t = h(x_t, u_t), \quad u_t \in U \\ & x_0 \in X, \quad \{x_t\}_{t \geq 0} \subset X \end{cases} \quad (2.7)$$

with a discount factor $0 < \delta \leq 1$ and a discount rate $\rho \geq 0$. The case where $\delta = 1$ and $\rho = 0$ is referred to as the undiscounted version of the problem (2.6) and (2.7). The asymptotic behaviour of the optimal path in the discounted framework depends critically on the magnitude of the initial stock and the discount factor.

Two classes of turnpike theorems should be distinguished when the future is discounted. The first class is the *asymptotic turnpike theorem*, see e.g. in Scheinkman [85], Bewley [6], Yano [87], Carlson [15], Carlson, Haurie and Jabrane [14], Zaslavski [89], Park [79]. It states that there exists a discount factor $\delta' \in (0, 1)$ such that for $x_0 \in X$ and any $\delta \in [\delta', 1)$, the optimal path $\{x_t^*\}$ starting at x_0 converges to the stationary path x^* .

The asymptotic turnpike theorem was proved under strictly concavity assumptions on L by e.g. Scheinkman [85] and Bewley [6], Majumdar, Nermuth [63], Kamihigashi and Roy [47]. Majumdar and Mitra [61] interpreted the function L as the consumption and set $L(x_t, x_{t+1}) := f(x_t) - x_{t+1} > 0$ for all t and considered a nonconvex (convex-concave) technology function f , i.e. f is assumed to be strictly increasing and twice continuously differentiable such that f satisfies $f'(\infty) < 1 < f'(0) < \infty$ and there is a real number $k_1 > 0$ such that $f''(x) = 0$ for $x = k_1$, $f''(x) > 0$ for $0 \leq x < k_1$ and $f''(x) < 0$ for $x > k_1$.

Montrucchio [72, 73] assumed certain curvature restrictions, i.e. the function L is continuous and strongly (α, β) -concave, i.e. $L(x, y) + \frac{1}{2}\alpha|x|^2 +$

$\frac{1}{2}\beta|y|^2$ is concave with $\alpha + \beta > 0$ and that the value function $W(x_0) := \max \sum_{t=0}^{\infty} \delta^t L(x_t, x_{t+1})$ is concave- γ for all $\delta \in (\delta_0, 1)$, i.e. $W(x) + \frac{1}{2}\gamma|x|^2$ is convex for all $\delta \in (\delta_0, 1)$. He proved that there exists $\delta' = 1 - \frac{\alpha+\beta}{\gamma}$ such that for a discount factor $\delta \in (\max[\delta_0, 1 - \frac{\alpha+\beta}{\gamma}], 1)$ the local asymptotic turnpike theorem holds, where $\alpha + \beta$ is a measure of the lower curvature of L and γ is a measure of its upper curvature. Carlso and Haurie [16] proved a version of the asymptotic turnpike theorem for infinite horizon open-loop differential games with $\delta = 1$ under a “strict diagonal concavity condition” on L . An *uniform asymptotic turnpike theorem* was proved by Cuong LeVan and Lisa Morhaim [58] under similar assumptions, i.e. if the function $L(x, y)$ is a strictly concave function, increasing in x and decreasing in y , then for any initial point $x_0 \in X$, there exists a $\beta' \in (0, 1)$ such that for any $\beta \in [\beta', 1)$, the asymptotic turnpike property holds.

Araujo and Scheinkman [5], McKenzie [68], Dasgupta and McKenzie [21] and Hyun Park [79] applied the *implicit function theorem* to prove the asymptotic turnpike theorem by assuming a dominant diagonal blocks condition: for an infinite matrix D formed by $n \times n$ blocks D_{ij} , $i, j \in \{1, 2, \dots\}$, D_{ii} is invertible, $\sup_i |D_{ii}^{-1}| < \infty$ and $\sup_i \sum_{j=1, i \neq j}^{\infty} |D_{ii}^{-1} D_{ij}| < 1$, where the norm $|D_{ij}| = \sup_{|z|=1} |D_{ij}z|$.

The second class is the *neighborhood turnpike theorem*, see e.g. [13, 69, 87, 18, 70, 28, 62, 47]. It states that for any $\epsilon > 0$, there exists $T' > 0$ and $\delta' \in (0, 1)$ such that, an optimal path starting from x_0 at the discount factor $\delta \in (\delta', 1)$ eventually stays within the ϵ -neighborhood of a stationary path x^* , i.e.

$$\|x_t^* - x^*\| \leq \epsilon, \quad \forall t \geq T'.$$

If a smaller neighborhood ϵ is chosen, the closer the discount factor δ is to 1.

To prove the neighborhood turnpike theorem, e.g. in McKenzie [70], Yano [87], instead of strictly concavity assumption, the differentiability of the function L is required. A version of *uniform neighborhood turnpike theorem* was proved in Guerrero-Luchtenberg and Leonardo [57]. It states that, under certain differentiability and concavity conditions on L , the neighborhood turnpike theorem holds for almost all initial states.

Finite-horizon problem with discounting was examined in e.g. [30, 28]. Based on convexity condition and using Implicit Programming formulation, Feinstein and Oren [30] proved a *funnel asymptotic turnpike theorem*, namely, an optimal path lies in an exponentially bounded region (determined by the discount factor) of the optimal stationary path. Fershtman and Kamien [28]

considered a finite-horizon differential games with quadratic cost function and proved a neighborhood turnpike theorem for the feedback equilibrium strategies.

2.2 The stochastic version of turnpike theorems

The turnpike theorems under uncertainty was first developed by William and Mirman [10]. The uncertainty or the stochastic environment is represented by a filtered probability space $(\Omega, \mathcal{F}, \nu; \{\mathcal{F}_t\}_{t \geq 0})$, $\omega \in \Omega$ is a possible state of the environment, \mathcal{F} is the Borel σ -field, ν is a probability measure on Ω , $\{\mathcal{F}_t\}_{t \geq 0}$ is the filtration satisfying the “usual” conditions, i.e. it is right-continuous and complete. The underlying evolution of the states $\{x_t\}$ is a stochastic process on $(\Omega, \mathcal{F}, \nu)$. Then the stochastic optimal control problem, in the undiscounted case, can be described, in the discrete-time setting as

$$\text{maximise } \mathbb{E} \left[\sum_{t=0}^T L(x_t, u_t) \right] \quad (2.8)$$

and in the continuous-time setting as

$$\text{maximise } \mathbb{E} \left[\int_0^T L(x_t, u_t) dt \right] \quad (2.9)$$

subject to the constraint $(x_t, u_t) \in X \times U$ with the initial condition $x_0 \in X$ and the terminal condition $x|_{t=T} = x_T \in X$ with probability 1. Here, in general, x_0 is an \mathcal{F}_0 -measurable random variable and x_T is an \mathcal{F}_T -measurable random variable. An optimal path (x_t^*, u_t^*) is an \mathcal{F}_t -adapted process such that (2.8) or (2.9) is achieved.

In the case where the terminal value is not given and a terminal payoff function S is considered, the problems can be reformulated as

$$\text{maximise } \mathbb{E} \left[\sum_{t=0}^T L(x_t, u_t) + S(x_T) \right] \quad (2.10)$$

and

$$\text{maximise } \mathbb{E} \left[\int_0^T L(x_t, u_t) dt + S(x_T) \right] \quad (2.11)$$

Further, when the future is discounted, the problem is described, in the discrete-time setting as

$$\text{maximise } \mathbb{E} \left[\sum_{t=0}^T \delta^t L(x_t, u_t) \right] \quad (2.12)$$

and in the continuous-time setting as

$$\text{maximise } \mathbb{E} \left[\int_0^T e^{-\rho t} L(x_t, u_t) dt \right]. \quad (2.13)$$

The stochastic analog of the turnpike theorem states that the optimal paths converge in a stochastic sense (say, convergence in distribution, convergence in probability or almost surely convergence) to a unique stochastic limit under standard assumptions of differentiability and concavity of the function L . The proof was mostly based on dynamic programming principle.

The most studied stochastic version of the middle turnpike theorem states that for any $\epsilon > 0$, there exists a number $\theta > 0$ such that the expected number of time periods for which $\{x_t^*\}$ spends outside the ϵ -neighbourhood of x^* , i.e. $\|x_t^* - x^*\| > \epsilon$, is bounded from above by θ , where θ is independent of T .

The uniformity assumption and the sensitivity assumption were mostly used to prove the stochastic version of turnpike theorems, as discussed in Joshi [46]. The idea is first to construct a value-loss process, say $\{V_t : t \geq 0\}$ which has the convenient property of being either a martingale (Majumdar and Zilcha [76]), a submartingale (Brock and Majumdar [12] or a supermartingale (Joshi [45]) thereby permitting a passage to the rich theory of martingales. Then the process $\{V_t : t \geq 0\}$ is assumed to be uniform bounded in expectation (i.e. the uniformity assumption) and have certain sensitivity property (i.e. the sensitivity assumption). The sensitivity assumption of $\{V_t : t \geq 0\}$ says that, for any $\epsilon > 0$, if the optimal paths diverge from the stochastic limit by more than ϵ , then the process $\{V_t\}$ records a value-loss of at least $\eta(\epsilon) > 0$. In other words, value-loss should be sensitive to a critical divergence of the optimal paths.

For general Markov chains the question on the existence of optimal value for discounted prices was addressed in [38]. Turnpike theorems for Markov chains were proved by Shapiro [83] in discounted case and by Kolokoltsov [49] without discounting. In the latter paper the turnpike theorems on the strategies and on the distributions of states were distinguished.

For the financial portfolio problems, the turnpike theorems was proved, mainly under convexity conditions on the function L directly, e.g. by Mossin [78], Leland [56], Hakansson [34], Hakansson [34], Sumit Joshi [45], Cox and Huang [19] and Huang and Zariphopoulou [40], Dybvig, Rogers and Back Kerry [24]. In the discrete-time setting, Mossin [78] showed the convergence of optimal strategies under the assumption that $-\frac{L'(x)}{L''(x)} = ax + b$, where L' and L'' is the first and second order derivative of L , respectively. Leland [56] generalised Mossin's result by assuming that $-\frac{L'(x)}{L''(x)} = ax + h(b)$ where the function h is uniformly bounded. Hakansson [34] proved convergence of optimal strategies by making assumptions slightly more general $\frac{(x-a)^{1-b}}{1-b} \leq L(x) \leq \frac{(x-a)^{1-b}}{1-b}, \forall x \geq a$. Huberman and Ross [41] showed that convergence occurs for the functions L that are bounded from below and for some $c \in (0, 1)$, $\lim_{x \rightarrow \infty} \frac{-L''(x)x}{L'(x)} = c$.

In the continuous-time setting, Cox and Huang [19] demonstrate that the portfolio turnpike property holds if the utility function L is strictly concave with $\lim_{x \rightarrow \infty} L'(x) = 0$ and there exists $A_1, A_2, b > 0$ and $K > 0$ such that $|(L')^{-1}(x) - A_1 x^{-\frac{1}{b}}| \leq A_2 x^{-a}$, for all $x \leq K$ and some $a \in [0, \frac{1}{b})$, where $(L')^{-1}$ denotes the inverse of L' . Huang and Zariphopoulou [40] gave another sufficient condition for the turnpike property that the strictly concave function L satisfies the condition $\lim_{x \rightarrow \infty} \frac{L'(x)}{x^r - 1} = k$ for $0 < r < 1$ and $k > 0$.

In optimal growth theory, mainly under the uniformity assumption and the sensitivity assumption, in the undiscounted case, Brock and Mirman [11], Dana [20], Evstigneev [25, 26, 4] and Mirman and Zilcha [75] showed that all optimal paths converge in an appropriate topology to the optimal stationary path. In the discounted context, Brock and Mirman [10] and Mirman and Zilcha [74] proved the convergence in distribution of optimal paths from distinct initial stocks to the stationary optimal paths. Fleming, Sethi and Soner [31] studied production planning problem where the demand is described as a Markov chain and showed a turnpike theorem for the Markov control.

Some research considered a so-called *non-stationary* setting. The stochastic environment can be non-stationary, namely, $\Omega := X_{t=0}^\infty \Omega_t$ and is considered as the set of all sequences $\omega := \{\omega_t : \omega_t \in \Omega_t, t \geq 0\}$, where ω_t is a possible state of the environment at time t . If the function L depends additionally on time t , it is called time-varying or non-stationary preference. In non-stationary models, the convergence in probability (Brock and Majumdar [12]) and the stronger property of almost sure convergence (Chang [17],

Föllmer and Majumda [29], Joshi [45], Majumdar and Zilcha [76], Joshi [46]) were proved. Denardo and Rothblum [22] considered an exponential utility function L for a Markov decision process and proved that if the corresponding transition rate matrix is transient, there exists an optimal stationary control path u^* .

In game theory, turnpike theorems for long-time two-player zero-sum stochastic games on a finite state space were proved by e.g. Kolokoltsov [50], Kolokoltsov and Malafeyev [51], by utilising the Bellman (or Shapley) operator and imposing conditions on the transition probability. Related research on long-time two-player zero-sum stochastic games can be founded e.g. in Sorin [84], Hernández-Lerma and Lasserre [37], Guo and Hernández-Lerma [33].

Haurie and Delft [36] considered a dynamics games with piecewise deterministic path and random jumps. In the discounted setting, under the assumption of convexity and piecewise uniqueness, the global asymptotic was proved if the transition probability q_{ij} are constant and the local asymptotic was proved if q_{ij} are state and control dependent.

For zero-sum Markov games, Jaśkiewicz [42], Vega-Amaya [86] worked with strongly continuous transition probabilities, i.e. the transition probability $q(D|x, u, v)$ is continuous in actions u and v , for any fixed states x and Borel subset D of the state space X . It was proved that the optimality equation has a solution, there exists a value of the game and both players possess optimal stationary strategies. In the proof, Jaśkiewicz [42] analyzes auxiliary perturbed models, whereas Vega-Amaya [86] makes use of a fixed point theorem.

In the work of Jaśkiewicz and Nowak [44] and Küenle [55], zero-sum Markov games within the similar setting where the transition times occur at integer time moments were studied. Jaśkiewicz and Nowak [44] prove the similar results by applying Fatou's Lemma for weakly convergent measures. Küenle [55] introduced certain contraction operators which lead to a parameterized family of functional equations. Making use of some continuity and monotonicity properties of the solutions to these equations (with respect to the parameter) he obtains a lower semicontinuous solution of the optimality equation. Then he showed that the maximizing player has an ϵ -optimal stationary strategy, and the minimizing player has an optimal stationary strategy.

Jaśkiewicz [43] studied a zero-sum ergodic semi-Markov games, in a general Borel state space, with weakly continuous (Feller) transition probabil-

ities and lower semicontinuous payoff function L . Under the Banach fixed point theorem and an ergodicity assumption of an embedded Markov chain, it was proved that in this game, one player possesses an ϵ -optimal stationary strategy and the other has an optimal stationary strategy.

Beyond the uniqueness assumption on the turnpike, it is worth noting that if there are several optimal stationary paths, there exists a state transition of the optimal paths from one optimal stationary path to another, see e.g. Mamedov [60], Alain Rapaport and Pierre Cartigny [81], Kamihigashi and Roy [47].

3 Basic setting for a zero-sum Markov game

In this section, we shall set up a model for a discrete-time two-person zero-sum Markov game on a general Borel state space. A Bellman (or Shapley) operator is defined for this game and some basic properties of this operator are discussed.

A discrete-time two-person zero-sum T -step, $T \in \mathbb{N}$, Markov game is a tuple (X, π, U, V, P, g, S) of the following meanings:

- X is a Borel subset of a general complete metric state space endowed with its Borel σ -field $\mathcal{B}(X)$ and a Borel σ -finite reference measure π on it. One can always keep in mind the sets $\{1, \dots, n\}$ and \mathbb{N} with arbitrary discrete measure or a subset of \mathbb{R}^d with Lebesgue measure as basic examples of X with its reference measure π .
- U and V are compact spaces of the admissible strategies (i.e., control parameters) for player 1 and player 2, respectively.
- $P = \{P(u, v, x, \cdot), x \in X, u \in U, v \in V\}$ is the family of transition probability kernels, i.e. $P(u, v, x, A)$, $A \in \mathcal{B}(X)$, is the transition probability from state x to A if the two players choose strategies u and v respectively. We assume that the transition probability $P(u, v, x, \cdot)$ has a density $\rho(u, v, x, y)$ with respect to the reference measure $\pi(\cdot)$.
- $g : U \times V \times X \times X \mapsto \mathbb{R}$ is a transition cost function for player 1, namely, $g(u, v, x, y)$ is the cost of the player 1 (i.e., the income of the player 2) from the transition from x to y under controls u and v . If the cost $g(u, v, x, y)$ is negative, this means that the player 2 pays $|g(u, v, x, y)|$ to player 1. The function g is assumed to be continuous and also bounded by certain constant $G > 0$, uniformly in u, v, x, y , i.e.

$$|g(u, v, x, y)| \leq G, \quad \forall u, v, x, y.$$

- $S : X \mapsto \mathbb{R}$ is a final cost function depending on the final position. The function S is assumed to be bounded and continuous. We denote the set of bounded and continuous functions on X by $C_b(X)$.

The dynamic T -step game with a given initial position $x \in X$ is played as follows. At the first step, starting from the given initial state x , the players choose sequentially certain strategies $v \in V$ and then $u \in U$ and the game moves to the next state $y \in X$ with probability $\rho(u, v, x, y)\pi(dy)$. Then the player 1 pays $g(u, v, x, y)$ to player 2 if $g(u, v, x, y)$ is positive; otherwise, the player 2 pays $|g(u, v, x, y)|$ to player 1. The second step is played analogously starting from the new position y . After T steps, the player 2 receives additional payment $S(k)$ depending on the final position $k \in X$. In this multi-step game, the player 1 aims to minimise the payoff to player 2.

We define the Bellman operator $B : C_b(X) \mapsto C_b(X)$ by

$$(BS)(x) := \inf_u \sup_v \left[\int_X (g(u, v, x, y) + S(y)) \rho(u, v, x, y) \pi(dy) \right]. \quad (3.1)$$

Then, according to the dynamic programming principle, the solution to the T -step Markov game is given by the iterations $B^T S$. We would like to mention that, since the transition cost function g and the final cost function S are assumed to be bounded and continuous, the inf sup in the definition (3.2) can be replaced by min max. Namely, in our setting, the Bellman operator B can be written as

$$(BS)(x) = \min_u \max_v \left[\int_X (g(u, v, x, y) + S(y)) \rho(u, v, x, y) \pi(dy) \right]. \quad (3.2)$$

Remark 3.1. *It is often assumed that the game defined by (3.2) has a value, which means that inf sup are interchangeable for all $S \in C_b(X)$ such that one also has*

$$(BS)(x) = \sup_v \inf_u \left[\int_X (g(u, v, x, y) + S(y)) \rho(u, v, x, y) \pi(dy) \right].$$

But we shall not make or use this assumption. Hence, in our framework, B^T describes the payoff for the player 1 calculated for the worst scenario dynamics (sometimes called robust control).

It is clear that, for any constant $a \in \mathbb{R}$ and any function $S \in C_b(X)$, the Bellman operator satisfies the following *homogeneity property*

$$B(a + S) = a + BS. \quad (3.3)$$

Moreover, the Bellman operator B is *non-expansive* in the usual sup-norm, i.e.

$$\|BS_1 - BS_2\| \leq \|S_1 - S_2\|, \quad \forall S_1, S_2 \in C_b(X) \quad (3.4)$$

since the inequality

$$\begin{aligned} \|BS_1 - BS_2\| &= \sup_{x \in X} |(BS_1 - BS_2)(x)| \\ &\leq \sup_{x \in X} \sup_{u,v} \int_X |(S_1(y) - S_2(y))| \rho(u, v, x, y) \pi(dy) \\ &\leq \|S_1 - S_2\| \end{aligned}$$

holds for any $S_1, S_2 \in C_b(X)$.

With the homogeneity property (3.3) and the non-expansive property (3.4) of the Bellman operator B , we can define the quotient space $\tilde{C}_b(X)$ of $C_b(X)$ with respect to constant functions. Namely, the equivalence classes are defined by

$$[S] := \{S + a, \forall a \in \mathbb{R} : S \in C_b(X)\}.$$

Let $\Pi : C_b(X) \mapsto \tilde{C}_b(X)$ be the natural projection. The quotient norm on $\tilde{C}_b(X)$ is defined by the formula

$$\|\Pi(f)\| = \inf_{a \in \mathbb{R}} \|f + a\| = \frac{1}{2} \left(\sup_{x \in X} f(x) - \inf_{x \in X} f(x) \right). \quad (3.5)$$

It is clear that Π has a unique isometric right inverse mapping $\Phi : \tilde{C}_b(X) \mapsto C_b(X)$ so that $\Pi \circ \Phi$ is an identity in $\tilde{C}_b(X)$. The image $\Phi(\tilde{C}_b(X))$ consists of all functions $S \in C_b(X)$ such that

$$\sup_{x \in X} S(x) = - \inf_{x \in X} S(x).$$

Thus one can identify $\tilde{S} \in \tilde{C}_b(X)$ with its image $S = \Phi(\tilde{S}) \in C_b(X)$. By the properties (3.3) and (3.4), the continuous quotient map $\tilde{B} : \tilde{C}_b(X) \mapsto \tilde{C}_b(X)$ is well defined.

4 Main results

In this section, we shall present the turnpike theorems for the game specified above. First, the uniqueness of the average cost and the stationary strategy

are proved. Then, the turnpike theorems on the set of strategies and on the state space are proved.

To state the main results, we need some additional properties of the transition probabilities:

$$\begin{aligned} \exists \delta > 0 : \exists A \in \mathcal{B}(X) \text{ with } \pi(A) > 0 : \forall x \in X, \forall u \in U, \forall v \in V, \forall y \in A, \\ \rho(u, v, x, y) \geq \delta. \end{aligned} \quad (4.1)$$

First, we show that the operator \tilde{B} is a contraction in the following

Lemma 4.1. *If the property (4.1) holds, then*

$$\|\tilde{B}(\tilde{S}_1) - \tilde{B}(\tilde{S}_2)\| \leq (1 - \delta\pi(A))\|\tilde{S}_1 - \tilde{S}_2\| \quad \forall \tilde{S}_1, \tilde{S}_2 \in \tilde{C}_b(X). \quad (4.2)$$

Proof. For arbitrary $S_1, S_2 \in C_b(X)$ and arbitrary states $x, z \in X$,

$$\begin{aligned} & (BS_1 - BS_2)(x) - (BS_1 - BS_2)(z) \\ & \leq \max_{u,v} \int_X (S_1(y) - S_2(y))\rho(u, v, x, y)\pi(dy) \\ & \quad - \min_{u,v} \int_X (S_1(y) - S_2(y))\rho(u, v, z, y)\pi(dy). \end{aligned}$$

Denote

$$(u_1, v_1) = \arg \max_{u,v} \int_X (S_1(y) - S_2(y))\rho(u, v, x, y)\pi(dy)$$

and

$$(u_2, v_2) = \arg \min_{u,v} \int_X (S_1(y) - S_2(y))\rho(u, v, z, y)\pi(dy).$$

Then, we have

$$\begin{aligned} & (BS_1 - BS_2)(x) - (BS_1 - BS_2)(z) \\ & \leq \int_X (S_1(y) - S_2(y)) \left(\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y) \right) \pi(dy) \\ & \leq \|S_1 - S_2\| \int_X |\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y)| \pi(dy). \end{aligned} \quad (4.3)$$

By property (4.1), for all $y \in A$,

$$\rho(u_1, v_1, x, y) \geq \delta, \quad \text{and} \quad \rho(u_2, v_2, z, y) \geq \delta.$$

Then

$$\begin{aligned} & \int_X |\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y)| \pi(dy) \\ & \leq \int_{X \setminus A} (\rho(u_1, v_1, x, y) + \rho(u_2, v_2, z, y)) \pi(dy) \\ & \quad + \int_A |\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y)| \pi(dy) \\ & = 1 - \int_A \rho(u_1, v_1, x, y) \pi(dy) + 1 - \int_A \rho(u_2, v_2, z, y) \pi(dy) \\ & \quad + \int_A |\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y)| \pi(dy). \end{aligned}$$

Denote by A^+ the set of $y \in A$ such that

$$\rho(u_1, v_1, x, y) > \rho(u_2, v_2, z, y)$$

and by A^- the set of $y \in A$ such that

$$\rho(u_1, v_1, x, y) \leq \rho(u_2, v_2, z, y).$$

Clearly, $A^+ \cap A^- = \emptyset$ and $A^+ \cup A^- = A$, so

$$\begin{aligned} & \int_X |\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y)| \pi(dy) \\ & \leq 2 - \int_{A^+ \cup A^-} \rho(u_1, v_1, x, y) \pi(dy) - \int_{A^+ \cup A^-} \rho(u_2, v_2, z, y) \pi(dy) \\ & \quad + \int_{A^+} (\rho(u_1, v_1, x, y) - \rho(u_2, v_2, z, y)) \pi(dy) \\ & \quad + \int_{A^-} (\rho(u_2, v_2, z, y) - \rho(u_1, v_1, x, y)) \pi(dy) \\ & = 2 - 2 \int_{A^+} \rho(u_2, v_2, z, y) \pi(dy) - 2 \int_{A^-} \rho(u_1, v_1, x, y) \pi(dy) \\ & \leq 2 - 2\delta\pi(A^+) - 2\delta\pi(A^-) \\ & = 2 - 2\delta\pi(A). \end{aligned}$$

Then, together with (4.3), we get

$$(BS_1 - BS_2)(x) - (BS_1 - BS_2)(z) \leq 2(1 - \delta\pi(A))\|S_1 - S_2\|.$$

Now by the definition of the quotient norm (3.5)

$$\|\Pi(BS_1 - BS_2)\| \leq (1 - \delta\pi(A))\|S_1 - S_2\|.$$

So we obtain

$$\|\tilde{B}\tilde{S}_1 - \tilde{B}\tilde{S}_2\| \leq (1 - \delta\pi(A))\|\tilde{S}_1 - \tilde{S}_2\|.$$

□

Theorem 4.2 (On the average cost). *If (4.1) holds, then there exists a unique $\lambda \in \mathbb{R}$ and a function $S^* \in C_b(X)$, unique up to equivalence (i.e. $\tilde{S}^* = \Pi(S^*)$ is unique), such that*

$$B(S^*) = \lambda + S^*. \quad (4.4)$$

Moreover, for every $S \in C_b(X)$,

$$\|B^T S - T\lambda\| \leq \|S^*\| + \|S^* - S\| \quad (4.5)$$

$$\lim_{T \rightarrow \infty} \frac{B^T S}{T} = \lambda \quad (4.6)$$

and, for every $\tilde{S} \in \tilde{C}_b(X)$,

$$\|\tilde{B}^T \tilde{S} - \tilde{S}^*\| \leq (1 - \delta\pi(A))^T \|\tilde{S} - \tilde{S}^*\|. \quad (4.7)$$

Proof. First, by the contraction property (4.2) of the operator \tilde{B} on $\tilde{C}_b(X)$ and the contraction principle (i.e. Banach fixed point theorem), the operator \tilde{B} has a unique fixed point in $\tilde{C}_b(X)$. Namely, there exists a unique $\tilde{S}^* \in \tilde{C}_b(X)$ such that

$$\tilde{B}\tilde{S}^* = \tilde{S}^*. \quad (4.8)$$

Consequently, there exists a $\lambda \in \mathbb{R}$ and a function $S^* \in C_b(X)$, unique up to equivalence such that (4.4) holds.

Next, we shall prove (4.5). On one hand, by (4.4) and (3.3)

$$B^T S^* = B^{T-1}(S^* + \lambda) = B^{T-2}(S^* + 2\lambda) = \dots = T\lambda + S^*$$

and together with the triangle inequality, for any $S \in C_b(X)$, we get

$$\|B^T S - B^T S^*\| = \|B^T S - T\lambda - S^*\| \geq \|B^T S - T\lambda\| - \|S^*\|. \quad (4.9)$$

On the other hand, by (3.4), we have,

$$\|B^T S - B^T S^*\| \leq \|B^{T-1} S - B^{T-1} S^*\| \leq \dots \leq \|S - S^*\|. \quad (4.10)$$

By combining (4.9) and (4.10), we have

$$\|B^T S - T\lambda\| - \|S^*\| \leq \|S - S^*\|$$

which clearly gives the inequality (4.5).

Equation (4.6) is implied by (4.5) by passing to the limit $T \rightarrow \infty$,

$$0 \leq \lim_{T \rightarrow \infty} \frac{\|B^T S - T\lambda\|}{T} \leq \lim_{T \rightarrow \infty} \frac{\|S^*\| + \|S - S^*\|}{T} = 0$$

as the functions S and S^* are all bounded functions. Equation (4.6), in turn, implies the uniqueness of λ in (4.4).

Further, for every $\tilde{S} \in \tilde{C}_b(X)$, by (4.8) we have

$$\begin{aligned} \|\tilde{B}^T \tilde{S} - \tilde{S}^*\| &= \|\tilde{B}(\tilde{B}^{T-1} \tilde{S}) - \tilde{B}(\tilde{B}^{T-1} \tilde{S}^*)\| \\ &\leq (1 - \delta\pi(A)) \|\tilde{B}^{T-1} \tilde{S} - \tilde{B}^{T-1} \tilde{S}^*\| \\ &\leq \dots \\ &\leq (1 - \delta\pi(A))^T \|\tilde{S} - \tilde{S}^*\| \end{aligned}$$

which completes the proof. \square

Remark 4.3. *The unique λ in (4.4) can be interpreted as the average cost of player 1 from the long time Markov game.*

Now we are ready for the first turnpike theorem on the set of optimal strategies. Here we introduce some notations. Denote by $E(x, S) = \{(u(x), v(x))\}$ the set of pairs of optimal (equilibrium) strategies $(u(x), v(x))$ of (3.2) for any $S \in C_b(X)$. Let $E(S) = \cup_{x \in X} \{E(x, S)\}$ be the set of pairs of optimal (equilibrium) strategies with final cost function $S \in C_b(X)$. In particular, the family $E(S^*)$ contains the sets of the stationary strategies in the game and $E(B^{T-t} S)$ is that the family of the sets of optimal (equilibrium) strategies at the step t of a T -step game with final cost function $S \in C_b(X)$.

Theorem 4.4 (Early turnpikes on the set of strategies). *Let (4.1) hold. Assume the set-valued function $E(\cdot)$ is continuous at least at the point $S =$*

S^* . Then for an arbitrary neighbourhood $U(E(S^*))$ of the set $E(S^*)$, there exists an $M \in \mathbb{N}$ such that if $T > M$, then for any $S \in C_b(X)$

$$E(B^{T-t}S) \subset U(E(S^*))$$

for all $t < T - M$.

Remark 4.5. The continuity of the set-valued function $E(\cdot)$ is in the sense of Hausdorff metric. The property that $E(\cdot)$ is continuous at the point $S = S^*$ has been used and discussed in many places. For example, in [39], this property is called the sensitivity property for a control law.

Proof. By (4.7), we have for any fixed T

$$\tilde{B}^{T-t}\tilde{S} \rightarrow \tilde{S}^* \quad \text{for large enough } T - t.$$

By the assumption that $E(S)$ depends continuously on S around S^* , we have

$$E(\tilde{B}^{T-t}\tilde{S}) \rightarrow E(\tilde{S}^*) \quad \text{for large enough } T - t.$$

So there exists an $M \in \mathbb{N}$, such that

$$E(\tilde{B}^{T-t}\tilde{S}) \subset U(E(\tilde{S}^*)), \quad t < T - M$$

which completes the proof. \square

Remark 4.6. This theorem states that, in a sufficiently large T -step game with an arbitrary final cost function $S \in C_b(X)$, the optimal strategies at steps $t < T - M$ must be close to the stationary strategies $E(S^*)$ and diverge away from the turnpike only near to the end of the game. Thus this is a stochastic competitive control version of the early turnpike as defined in [67].

Further, with additional assumptions on the set of optimal strategies and on the equilibrium transitions, we shall show that, in a long-time game, if the game is carried out with optimal strategies, then, apart from short initial and final periods of time, the distributions of its position on the state space must be very close to the stationary distribution. Let $\|\cdot\|_{TV}$ denote the total variation norm of measures.

Theorem 4.7 (Middle turnpike on the (distribution of) state space). *Let (4.1) hold. Assume the set-valued function $E(S)$ is continuous at least at the point $S = S^*$ and, for each state $x \in X$, the set $E(x, S^*)$ contains only*

one pair of optimal strategies $(u^*(x), v^*(x))$. Let $Q^* = \{q^*(x)\}_{x \in X}$ denote the stationary distribution for the stationary Markov chain defined on the state space X by the unique pair of optimal strategies $(u^*(x), v^*(x))$. Assume also that the equilibrium transition probabilities depend locally Lipschitz continuous on S around S^* , i.e. that there exists a constant k such that,

$$\|P(u(x), v(x), x, \cdot) - P(u^*(x), v^*(x), x, \cdot)\|_{TV} \leq k\|S - S^*\|, \quad \forall x \in X \quad (4.11)$$

where $(u, v) \in E(x, S)$ and S is sufficiently close to S^* . Then for each $\epsilon > 0$, there exists an $M \in \mathbb{N}$ such that for each T -step game, $T > 2M$, with terminal cost $S \in C_b(X)$ of player 1, we have

$$\|Q(t) - Q^*\|_{TV} < \epsilon$$

for all $t \in [M, T - M]$, where $Q(t) = \{q(x, t)\}_{x \in X}$ and $q(x, t)$ is the probability that the process is in a state $x \in X$ at time t if the game is carried out with the optimal strategies.

Remark 4.8. In other words, $q^*(x)$ is the mean amount of time that each sufficiently long game with optimal strategies spends in position x .

Proof. From (4.11), together with Lemma 4.1, we have

$$\|P_t(x, \cdot) - P^*(x, \cdot)\|_{TV} \leq k\epsilon^{T-t}\|S - S^*\|$$

where $\epsilon = 1 - \delta\pi(A)$, $P_t(x, \cdot)$ is the transition probability arising from optimal strategies $(u(x), v(x))$ at time $t \leq T$ with terminal payoff S and $P^*(x, \cdot) = P(u^*(x), v^*(x), x, \cdot)$.

Let Ψ_t and Ψ^* denote linear operators on the set of measures

$$(\Psi_t \mu)(dx) := \int \mu(dz) P_t(z, dx) \quad \text{and} \quad (\Psi^* \mu)(dx) := \int \mu(dz) P^*(z, dx)$$

so that

$$\begin{aligned} \|\Psi_t - \Psi^*\| &= \sup_{\|\mu\|=1} \int \mu(dz) (P_t(z, dx) - P^*(z, dx)) \\ &\leq k\epsilon^{T-t}\|S - S^*\|. \end{aligned}$$

The state distribution at time t can be estimated by

$$\begin{aligned} Q(t) &= \Psi_t \cdots \Psi_1 Q(0) \\ &= (\Psi^* + \xi_t) \cdots (\Psi^* + \xi_1) Q(0) \end{aligned}$$

with $\xi_j \leq k\epsilon^{T-j}$, where $Q(0)$ is the initial state distribution. Since

$$\begin{aligned}
& \|\Psi_t \cdots \Psi_1 - (\Psi^*)^t\| \\
&= \|(\Psi_t - \Psi^*)\Psi_{t-1} \cdots \Psi_1 + \Psi^*(\Psi_{t-1} - \Psi^*)\Psi_{t-2} \cdots \Psi_1 \\
&\quad + \cdots + (\Psi^*)^{t-1}(\Psi_1 - \Psi^*)\| \\
&\leq k [\epsilon^{(T-t)} + \epsilon^{(T-t+1)} + \cdots + \epsilon^{(T-1)}] \\
&\leq k\epsilon^{(T-t)}(1 + \epsilon + \epsilon^2 + \cdots) \\
&= \frac{k}{1-\epsilon}\epsilon^{(T-t)} \leq \frac{k}{1-\epsilon}\epsilon^{M_1}
\end{aligned}$$

for $t < T - M_1$, $M_1 \in \mathbb{N}$, we have

$$\|Q(t) - (\Psi^*)^t Q(0)\|_{TV} \leq \epsilon_1$$

for $t < T - M_1$, where ϵ_1 can be made arbitrary small by choosing large enough $M_1 \in \mathbb{N}$.

By the theorem on the convergence of probability distribution in homogeneous Markov chains to a stationary distribution, it follows that for each $\epsilon_2 > 0$ there exists an $M_2 \in \mathbb{N}$ such that

$$\|(\Psi^*)^t Q(0) - Q^*\|_{TV} \leq \epsilon_2, \quad t > M_2.$$

Let $M = \max\{M_1, M_2\}$. If $T > 2M$, the intersection of the sets $[1, T - M_1]$ and $[M_2, \infty)$ is non-empty. Then we have

$$\|Q(t) - Q^*\|_{TV} \leq \epsilon_1 + \epsilon_2 \quad \forall t \in [M, T - M]$$

which completes the proof. \square

5 Concluding remarks

We have considered a two-person zero-sum Markov game with a general Borel state space and the modelling is formulated in the discrete time setting. With the Bellman operator defined by (3.2), we have proved the convergence of the long time average cost (reward) of the game and turnpike theorems on the set of optimal strategies and on the state space are proved. We would like to mention that these results still hold under the corresponding continuous time setting by discretising, that is by looking at a continuous time game as the limit of games with discrete times.

Notice that in our modelling, at each step, we can observe the exact position of this game. But it often happens that one has to take into account not only the exact position, but also, the state distribution (say, its variance of volatility). This leads to the turnpike properties for measure-valued Markov games, see [52, 53, 54]. For example, the final cost function may be any continuous function on a measure space and the stochastic transition kernels may depend on the state distribution. We shall develop results under this more general setting in future work.

References

- [1] Atsumi, H. (1965). Neoclassical Growth and the efficient program of capital accumulation. *Review of Economic Studies*, 32(2), 127-36.
- [2] Alvarez O., M. Bardi (2007). Ergodic problems in differential games. *Advances in Dynamic Game Theory*, S. Jorgensen, M. Quincampoix, and T.L. Vincent, eds., pp 131-152, Ann. Internat. Soc. Dynam. Games, vol. 9, Birkhaeuser, Boston.
- [3] Alvarez O., M. Bardi (2010). Ergodic Ergodicity, stablization, and singular perturbations of Bellman-Isaacs equations, *Memoirs American Math. Soc.*, 77 pages.
- [4] Arkin V. I., I.V. Evstigneer (1987). *Stochastic Models of Control and Economic Dynamics (Economic Theory, Econometrics, and Mathematical Economics)*. Academic Press.
- [5] Araujo, A. Scheinkman, J.A. (1977). Smoothness, comparative dynamics, and the turnpike property. *Econometrica*, 45(3), 601-620.
- [6] Bewley T.F. (1982). An integration of equilibrium theory and turnpike theory. *Journal of mathematical economics*, 20, 233-267.
- [7] Bewley, T.F.(2007). General equilibrium, Overlapping Generations Models and Optimal growth theory, Harvard University Press, Cambridge.
- [8] Bardi M. (2009). On differential games with long-time-average cost. *Annals of the International Society of Dynamic Games*, 10(1), 1-16.
- [9] Blot J., Crettez, B (2007). On the smoothness of optimal paths II: some local turnpike results. *Dicisions Econ Finan*, 30, 137-150.

- [10] Brock, W.A., Mirman, L.J. (1972). Optimal economic growth and uncertainty: the discounted case. *Journal of Economic Theory*, 4, 479-513.
- [11] Brock, W.A., Mirman, L.J. (1973). Optimal economic growth and uncertainty: the no discounting case. *International Economic Review*, 14(3), 560-573.
- [12] Brock, W.A., Majumdar, M. (1978). Global asymptotic stability results for multisector models of optimal growth under uncertainty when future utilities are discounted. *Journal of Economic Theory*, 18, 225-243.
- [13] Cass D. (1966). Optimum Growth in an Aggregative Model of Capital Accumulation: A Turnpike Theorem. *Econometrica*, 34(4), 833-850.
- [14] Carlson D.A.(1990). The existence of catching-up optimal solutions for a class of infinite horizon optimal control problems with time delay. *SIAM Journal on Control and Optimization*, 28, 402-422.
- [15] Carlson D.A., Jabrane A., Haurie A. (1987). Existence of overtaking solutions to infinite dimensional control problems on unbounded time intervals. *SIAM Journal on Control and Optimization*, 25, 1517-1541.
- [16] Carlson D.A., Haurie A.B. (1996). A turnpike theory for infinite horizon open-loop differential games with decoupled controls. *SIAM Journal on Control and Optimization*, 34(4), 1405-1419.
- [17] Chang, F.R. (1982). A note on the stochastic value loss assumption. *Journal of Economic Theory*, 26, 164-170.
- [18] Coles, J.L. (1985). Equilibrium turnpike theory with constant returns to scale and possibly heterogeneous discount factors. *International Economic Review*, 26(3), 671-679.
- [19] Cox J.C., C.-F. Huang (1992). A continuous-time portfolio turnpike theorem. *Journal of Economic Dynamics and Control*, 16, 491-507.
- [20] Dana, R.A.(1974). Evaluation of development programs in a stationary stochastic economy with bounded primary resources. In: Proceedings of the Warsaw Symposium on Mathematical Methods in Economics. North-Holland, Amsterdam, 179-205.
- [21] Dasgupta, S., Mckenzie, K.W. (1985). A note on comparative dynamics of stationary states. *Econ. Lett*, 18, 333-338.

- [22] Denardo E.V., Rothblum U.G. (2006). A turnpike theorem for a risk-sensitive Markov decision process with stopping. *SIAM J. Control Optimal*, 45(2), 414-431.
- [23] Dindo P. (2005). A tractable evolutionary model for the Minority Game with asymmetric payoffs. *Physica, A* (355), 110-118.
- [24] Dybvig P.H., Rogers L.C. G., Back Kerry (1999). Portfolio Turnpikes, the Review of Financial Studies, 12(1), 165-195.
- [25] Evstigneev I.V. (1974). Optimal stochastic programs and their stimulating prices. *mathematical Models in Economics*, J. Los and M.W. Los (eds.), 219-252, North-Holland.
- [26] Evstigneev I.V. (1976). Turnpike theorems in probabilistic models of economic dynamics. *Mathematical Notes* 19(2), 165-171, (translated from *Matematicheskie Zametki*, Vol. 19(2), 279-290, 1976).
- [27] Dorfman R., P. Samuelson and R. Solow (1958). *Linear programming and economic analysis*. New York: McGraw-Hill.
- [28] Fershtman C., M. I. Kamien (1990). Turnpike properties in a finite horizon differential game: dynamic duopoly with sticky prices. *International Economic Review*, 31(1), 49-60.
- [29] Föllmer, H., Majumdar, M. (1978). On the asymptotic behaviour of stochastic economic processes. *Journal of Mathematical Economics*, 5, 275-287.
- [30] Feinstein C.D., Oren S.S. (1985). A "funnel" turnpike theorem for optimal growth problems with discounting. *Journal of Economic Dynamics and Control*, 9, 25-39.
- [31] Fleming W., S.P. Sethi, Soner H.M. (1987). An optimal stochastic production planning problem with randomly fluctuating demand. *SIAM J. control Optim*, 25, 1494-1502.
- [32] Gale, D. (1967). On optimal development in a multi-sector economy. *Rev. Econ. Stud.*, 34, 1-18.
- [33] Guo X. and O. Hernández-Lerma (2005). Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates. *Bernoulli*, 11, 1009-1029.
- [34] Hankansson, N. (1974). Convergence to isoelastic utility and policy in multi-period portfolio choice. *Journal of Financial economics*, 1, 201-224.

- [35] Haurie A. (1976). Optimal control on an infinite time horizon: the turnpike approach. *Journal of mathematical economics*, 3(1), 81-102.
- [36] Haurie A., Delft C.V.(1991). Turnpike properties for a class of piecewise deterministic systems arising in manufacturing flow control. *Annals of operations Research*, 29, 351-374.
- [37] Hernandez-Lerma O. and J. B. Lasserre (2001). Zero-Sum Stochastic Games in Borel Spaces: Average Payoff Criteria, *SIAM J. Control Optim.*, 39, 1520-1539.
- [38] U. D. Holzbaaur. Fixed point theorems for discounted finite Markov decision processes. *J. Math. Anal. Appl.* 116:2 (1986), 594-597.
- [39] Huang M.Y., R.P. Malhame, P. E. Caines (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information and Systems*, 6(3), 221-251.
- [40] Huang C.Fu., Zariphopoulou T. (1999). Turnpike behaviour of long-term investments. *Finance Stochastic*, 3, 15-34.
- [41] Huberman G. and Ross S. (1983). Portfolio turnpike theorems, risk aversion and regularly varying functions. *Econometrica*, 51, 1345-1361.
- [42] Jaśkiewicz A. (2002). Zero-sum semi-Markov games. *SIAM J. Control Optim.*, 41, 723-739.
- [43] Jaśkiewicz A. (2009). Zero-sum ergodic semi-Markov games with weakly continuous transition probabilities. *J Optim Theory App*, 141, 321-347.
- [44] Jaśkiewicz A., Nowak, A.S. (2006). Zero-sum ergodic stochastic games with Feller transition probabilities. *SIAM J. control Optimal*, 45, 773-789.
- [45] Joshi S. (1997). Turnpike theorems in Nonconvex nonstationary Environments. *International econmic review*, 38(1), 225-248.
- [46] Joshi S. (2003). The stochastic turnpike property without uniformity in convex aggregate growth models. *Journal of economic Dynamics and control*, 27, 1289-1315.
- [47] Takashi K., Roy S. (2006). Dynamic optimization with a non smooth, non-convex technology: the case of a linear objective function. *Economic Theory*, 29, 325-340.

- [48] Khan M. A. and Piazza A. (2011). An overview of turnpike theory: towards the discounted deterministic case. *Adv. Math. Econ*, 14, 39-67.
- [49] Kolokoltsov V.N. Turnpikes and infinite extremals in Markov decision processes. (Russian) *Mat. Zametki* 46:4 (1989), 118120.
- [50] Kolokoltsov V.N. (1992). On linear, additive, and homogeneous operators in Idempotent Analysis. *Advances in Soviet Mathematics* 13, Idempotent Analysis, Ed. V.P.Maslov et S.N. Samborski, 87-101.
- [51] Kolokoltsov V.N. and O.A. Malafeyev (2010). Understanding Game Theory: Introduction to the analysis of many agent systems of competition and cooperation. World Scientific.
- [52] Kolokoltsov V.N. (2010). Nonlinear Markov Processes and Kinetic Equations. Cambridge University Press, Cambridge.
- [53] Kolokoltsov V.N. (2012). Nonlinear Markov games on a finite state space (mean-field and binary interactions). to appear in *International Journal of Statistics and Probability*.
- [54] Kolokoltsov V.N. , J. Li and W. Yang (2012). Mean Field Games and Nonlinear Markov Processes. arXiv:1112.3744v
- [55] Küenle, H.U.(2007). On Markov games with average reward criterion and weakly continuous transition probabilities. *SIAM J. Control Optim*, 45, 2156-2168.
- [56] Leland, H. (1972). On turnpike portfolios, In G. Szego and K. Shell, (eds.), Mathematical Methods in Investment and Finance. Amsterdam: North-holland.
- [57] Guerrero-Luchtenberg, César Leonardo (2000). A uniform neighborhood turnpike theorem and applications. *J. Nath. Econm.*, 34(3), 329-357.
- [58] Le Van C., Morhaim L.(2006). On optimal growth models when the discount factor is near 1 or equal to 1. *International Journal of economic theory*, 2(1), 55-76.
- [59] Mamedov M. A. (2002). Asymptotical stability of optimal paths in nonconvex problems, Research Report 02/02, University of Ballarat, February .
- [60] Mamedov, M. A. (2003). A turnpike theorem for continuous-time control systems when the optimal stationary point is not unique. *Abstract and Applied Analysis*, 11, 631-650.

- [61] Majumdar M., Mitra T. (1982). Intertemporal allocation with a Non-convex technology: the aggregative framework. *Journal of Economic theory*, 27, 101-136.
- [62] Marena, M., Montrucchio, L.(1999). Neighborhood turnpike theorem for continuous time optimization models. *J. Optim. Theory appl.*, 101, 651-676.
- [63] Majumdar and Nermuth (1982). Dynamic Optimisation in Non-Convex Models with Irreversible investment: Monotonicity and Turnpike Theorems. *Journal of Economics*, 42, 339-362
- [64] Makarov V.L. and A. M. Rubinov (1973). Mathematical Theory of Economic Dynamics and Equilibria, Nauka, Moscow, English trans.: Springer-Verlag, New York.
- [65] Maitra A. and Sudderth W. (1993). Borel stochastic games with limsup payoffs. *Ann. probab.*, 21, 861-885.
- [66] McKenzie L. W. (1963). The turnpike theorem of Morishima. *Review of Economic Studies*, 30, 169-176.
- [67] McKenzie L. W. (1976). Turnpike theory. *Econometrica*, 44(5), 841-865.
- [68] McKenzie L. W. (1977). A new route to the turnpike. In Henn, R. Moeachlin, O. (eds.) *Mathematical Economics and Games Theory*, Springer, New York
- [69] McKenzie, L. W. (1979). Optimal Economic growth and turnpike theorems. Discussion paper 79-1, university of Rochester.
- [70] McKenzie L. W. (1986). Optimal economic growth, and turnpike theorems and comparative dynamics, In: *Handbook of Mathematical Economics*, 3, 1281-1355.
- [71] McKenzie, L.W.(1998). Turnpikes. *The American Economic Review, Papers and Proceedings of the Hundred and Tenth Annual meeting of the American Economic Association*. 88(2), 1-14.
- [72] Montrucchio L. (1995). A turnpike theorem for continuous-time optimal-control models. *Journal of economic Dynamics and Control*, 19, 599-619.
- [73] Montrucchio L. (1995). A new turnpike theorem for discounted programs. *Econ. theory*, 5, 371-382.
- [74] Mirman, L.J., Zilcha, I. (1975). On optimal growth under uncertainty. *Journal of Economic Theory*, 11, 329-339.

- [75] Mirman, L.J., Zilcha, I. (1977). Characterizing optimal policies in a one-sector model of economic growth under uncertainty. *Journal of Economic Theory*, 14, 389-401.
- [76] Majumdar, M., Zilcha, I. (1987). Optimal growth in a stochastic environment: some sensitivity and turnpike results. *Journal of Economic Theory*, 43, 116-133.
- [77] Morishima, Michio (1961). Proof of a Turnpike theorem: the no joint production case, *Review of Economic Studies*, 28, 89-97.
- [78] Mossin, J.(1968). Optimal multiperiod policies. *J. Business*, 31, 215-229.
- [79] Park H. (2000). Global asymptotic stability of a competitive equilibrium with recursive preferences. *Economic theory*, 15(3), 565-584.
- [80] Radner, Roy (1961). Paths of Economic growth that are optimal with regard only to Final States. *Review of Economic Studies*, 28, 98-104.
- [81] Alain Rapaport and Pierre Cartigny (2004). Turnpike theorems by a value function approach. *SAIM: control, Optimisation and Calculus of Variations*, 10, 123-141.
- [82] Samuelson, P.A., R.M. Solow (1956). A complete capital model involving heterogeneous capital goods. *Q. J. Econ*, 27, 537-562.
- [83] J. F. Shapiro. Turnpike Planning Horizon for a Markov Decision Model. *Management Science* 14 (1968), 292-300.
- [84] Sorin S. (2005). New approaches and recent advances in two-person zero-sum repeated games. *Advances in Dynamic Games*, Nowak, Andrzej S.; Szajowski, Krzysztof (Eds.), 67-94,
- [85] Scheinkman J.(1976). On optimal steady states of n -sector growth models when utility is discounted. *Journal of Economic theory*, 12, 11-30.
- [86] Vega-Amaya, O. (2003). Zero-sum semi-Markov games. fixed-point solutions of the Shapley equation. *SIAM J. control Optm*, 42, 1876-1894.
- [87] Yano M.(1984). Competitive equilibria on turnpikes in a McKenzie Economy, I: A neighborhood turnpike theorem, *International Economic review*, 25(3) 695-717.
- [88] Yakovenko S.Yu., L.A. Kontorer (1992). Nonlinear semigroups and infinite horizon optimization, *Advances in Soviet mathematics*, 13,167-210.

- [89] Zaslavski A.J. (1995). Optimal programs on infinite horizon 1. *SIAM Journal on Control and Optimization*, 33, 1643-1660.
- [90] Zaslavski A.J. (1995). Optimal programs on infinite horizon 2, *SIAM Journal on Control and Optimization*, 33, 1661-1686.
- [91] Zaslavski A.J. (1999). The turnpike property for dynamic discrete time zero-sum games. *Abstract Applied Analysis*, 4(1), 21-48.
- [92] Zaslavski A.J. (1998). Turnpike theorem for convex infinite dimensional. *Journal of convex analysis*, 5(2), 237-248.
- [93] Zaslavski A.J. (2004). Turnpike theorem for a class of discrete time optimal control problem. *Fixed point theory and applications*, 5, 175-182.
- [94] Zaslavski A.J. (2006). Turnpike Properties in the Calculus of Variations and Optimal Control, Springer, New York.
- [95] Zaslavski A.J. (2007). Turnpike results for a discrete-time optimal control system arising in economic dynamics. *Nonlinear Analysis*, 67, 2024-2049.
- [96] Zaslavski A.J. (2009). Two turnpike results for a discrete-time optimal control system. *Nonlinear Analysis*, 71, 902-909.